ESL-P-492

THE INFIMUM PRINCIPLE

by

Hans P. Geering*) and Michael Athans**)

Abstract

This paper reports a reasonably complete theory of necessary and sufficient conditions for a control to be superior with respect to a non-scalar-valued performance criterion. The latter maps into a finite-dimensional integrally closed directed partially ordered linear space. The applicability of the theory to the analysis of dynamic vector estimation problems and to a class of uncertain optimal control problems is demonstrated.

The research presented in this document was made possible through the support extended by the U. S. Air Force Office of Scientific Research under grants AFOSR-70-1941 and AFOSR-69-1724 (supplement A) and by NASA under grant NGL-22-009-(124) to the Decision and Control Sciences Group of the M.I.T. Electronic Systems Laboratory.

*) Oerlikon-Buehrle AG, 8050 Zurich, Switzerland

**) Massachusetts Institute of Technology, Department of Electrical Engineering, Cambridge, Massachusetts 02139

I. INTRODUCTION

For optimal control and estimation problems, the Pontryagin minimum principle [1] allows finding the optimal parameters of a system (controls, gains, etc.), provided the performance criterion is a scalar-valued function of the parameters and the state variables.

The user of this optimization technique often experiences difficulties in choosing a suitable scalar-valued optimality criterion, because he may be interested in several cost functionals simultaneously, such as energy consumption, transfer time, and integral squared error, or because in vector estimation problems the estimation error is described more completely by the mean error vector and the error covariance matrix than by the trace or the determinant of the latter.

It then is natural to consider non-scalar-valued performance criteria, e.g. vectors, matrices, sets, etc. The meaning of "better than" has to be defined, which mathematically speaking is done by introducing a partial order relation [4], [19], [2] in the space of the cost functional. For non-scalar-valued performance criteria the notion of optimality splits up into the more restrictive superiority and the weaker non-inferiority (Pareto optimality), because "better than" is not the complement of "worse than".

In [21], vector-valued performance criteria were introduced. In this case, a superior solution minimizes all of the components of the cost vector simultaneously, whereas for noninferior solutions, no other solution can be found, which simultaneously improves all of the components of the cost vector. In [6],[7], and [8], a theory of non=inferior solutions to optimal control problems with vector-valued performance criteria has been developed. In [14], this theory has been extended to abstract partially ordered range spaces of the cost functional.

The problem of finding the more interesting superior solutions, called infimization problem, has not been considered in the control literature, mainly on the ground that one usually has contradicting cost objectives anyway. However, Ritter [16] solved infimization problems of mathematical programming in abstract infinite-dimensional partially ordered spaces with equality and inequality type of constraints.

The goal of this research was finding necessary conditions in the form of an infimum principle for a control to be a superior solution to a dynamic infimization problem, where the cost criterion takes its values in a suitable finite-dimensional partially ordered space. Some of the results, which were obtained, were reported in [9] and [10].

The purpose of this paper is to publish a reasonably complete theory of superior solutions to optimal control problems with non-scalar-valued performance criteria, together with some nontrivial applications. This theory is particularly useful in the analysis of optimal least-squares vector estimation problems (matrix-valued cost) and uncertain optimization problems with set membership description of uncertainty (set-valued cost).

The most pertinent definitions and facts of the theory of partial order relations can be found in the paper "Necessary and Sufficient Conditions for Differentiable Non-Scalar-Valued Functions to Attain Extrema" [2], and they are widely used in this paper. Some additional facts are collected in the Appendix.

Sections II and III contain the theoretic part of this paper. In Section II, two optimal control problems are stated. Necessary conditions for a control to be a superior solution are developed for both of the problems. These new results are discussed and compared with the Pontryagin minimum principle [1] and with the necessary conditions for noninferiority (which is less restrictive than superiority is) [6], [7], [8], [14].

In Section III, some of the sufficiency conditions for optimality for scalar-valued cost functionals are extended to sufficiency conditions for superiority for non-scalar-valued costs.

Section IV contains some nontrivial applications. The Kalman-Bucy filter is shown to be the superior solution to an optimization problem with matrix-valued cost (partial order of positive-semidefinite differences) [10]. The dual problem to the filtering problem, viz. the linear-quadratic regulator problem is also discussed. Furthermore, an uncertain optimization problem with set membership description of the uncertainty is shown to have a superior solution (partial order of set inclusion).

The contributions of this paper to the state of the art of optimal control theory are discussed in the concluding Section V.

II. NECESSARY CONDITIONS FOR SUPERIORITY

In the statement of the optimal control problems and in the proof of the theorems, this section basically follows the approach taken by Halkin [11] for optimal control problems with scalar-valued cost functionals. Thus, the cost space is a k-dimensional partially ordered subspace of the (extended) n-dimensional state space.

The statements of two dynamic infimization problems (superiority) are given in Section II.1. Section II.2 contains the results of these optimization problems in the form of necessary conditions for a control to be superior, which are proved in Section II.3. These results are new. A preliminary report has been given by the authors in [10]. However, Ritter has investigated a related problem of mathematical programming in [16]. The solution to the dynamic minimization problem (noninferiority) is essentially well known [6], [7], [8], but only Neustadt [14] considered abstract partially ordered cost spaces.

In Section II.4, the results of Section II.2 are discussed and their relation to [6], [7], [8], and [14], and to Pontryagin's minimum principle [1] is explained.

II.1 Statement of the Infimization Problems

In the statement of the optimal control problems, the cost space is a k-dimensional partially ordered subspace of the (extended) n-dimensional state space. Following Halkin [11], the dynamics of the state are allowed to depend on all of the n state variables. The cost space always is a finite-dimensional integrally closed directed linear poset [2, Definitions II.7, II.9, and II.10]. Hence, by Theorem II.13 of [2], its positive cone is closed

and has non-empty interior, and the complement of the positive cone is an open subset of the cost space. All finite-dimensional spaces involved are Banach spaces, but not neces= sarily Euclidean spaces. This allows treatment of states, controls, and costs to be vectors, matrices, etc. without use of canonical transformations. All derivatives are Fréchet deri= vatives, all measures are Lebesgue measures.

Given the dynamical system

$$\dot{x}(t) = f[x(t),u(t),t] \qquad \text{a. e. } t \in [t_0,t_1] \tag{1}$$

where $x(t) \in X^n$ is the state, $u(t) \in \Omega \subset X^m$ is the control, and $t \in [t_0,t_1] \subset R$ is the time, with fixed initial state x_0 at the fixed initial time t_0,

$$x(t_0) = x_0 \ . \tag{2}$$

The final time t_1 is fixed.

Assume, there exists $\varepsilon > 0$, such that $f[x,u,t]$ and $\partial f[x,u,t]/\partial x$ are defined, meas= urable with respect to u and t, uniformly equicontinuous with respect to x, and uniformly bounded for all $(x,t,u) \in N(x,t,\varepsilon) \times \Omega^*$, where $N(x,t,\varepsilon)$ is any closed ε-neighbourhood of $(x,t) \in X^n \times [t_0,t_1]$ and Ω^* any closed and bounded subset of Ω.

The admissible control functions u are such that $u(t) \in \Omega$ for all $t \in [t_0,t_1]$ and that there exists a state trajectory $x : [t_0,t_1] \to X^n$ satisfying (2) (which is then granted to be unique [11]).

The k-dimensional cost space $(X^k,\preccurlyeq)$ is an integrally closed directed linear poset and a subspace of the state space X^n. Thus, the state can be written as

$$x(t) = (S(t),J(t)) \tag{3}$$

where $J(t)$ is the component of the state, which belongs to the cost space and $S(t)$ the component, which belongs to the orthogonal complement $(X^k,\preccurlyeq)^{\perp}$ of the cost space (in the sense of direct sum).

In this section, the following two infimization problems are considered.

Infimization Problem II.1. Find an admissible control u^*, such that the final state satisfies the constraint

$$S(t_1;u^*) = S_1 \quad , \tag{4}$$

where S_1 is a fixed element of $(X^k,\preccurlyeq)^\perp \subset X^n$, and such that the cost component $J(t_1;u^*)$ of the optimal final state $x(t_1;u^*)$ is infimized, i.e.,

$$J(t_1;u^*) \preccurlyeq J(t_1;u) \tag{5}$$

for all admissible controls u meeting the boundary constraint, $S(t_1;u) = S_1$.

Infimization Problem II.2. Find an admissible control u^*, such that the cost component of the final state is infimized, i.e.,

$$J(t_1;u^*) \preccurlyeq J(t_1;u) \text{ for all admissible } u. \tag{6}$$

The Problem II.1 is a fixed-end-point problem whereas the Problem II.2 is a free-end-point problem.

II.2. Necessary Conditions for Superiority

Convention II.3. $P(t) \in \mathcal{L}(X^n,(X^k,\preccurlyeq))$ denotes a linear map from the state space X^n into the cost subspace $(X^k,\preccurlyeq)$. Furthermore, $P(t)$ is decomposed into $P(t) = (P_S(t),P_J(t))$, where $P_S(t) \in \mathcal{L}((X^k,\preccurlyeq)^\perp,(X^k,\preccurlyeq))$ and $P_J(t) \in \mathcal{L}((X^k,\preccurlyeq),(X^k,\preccurlyeq))$.

Definition II.4. For all $t \in [t_0,t_1]$, the Hamiltonian is defined by

$$H[x(t),P(t),u(t),t] = P(t)f[x(t),u(t),t] \tag{7}$$

where $f[x(t),u(t),t]$ is the right-hand side of (1). Note, that H attains its value in the k-dimensional cost subspace $(X^k,\preccurlyeq)$ of X^n.

Theorem II.5. Solution to Infimization Problem II.1

In order that the admissible control u^* be a superior solution to the Infimization Problem II.1, it is necessary, that there exist a nonzero map $(P_S, P_0) \in \mathcal{L}(X^n, (X^k, \preccurlyeq))$, where P_0 is positive (see Definition A.1), i.e., $P_0 \succcurlyeq 0 \in \mathcal{L}((X^k, \preccurlyeq), (X^k, \preccurlyeq), \preccurlyeq)$, such that along the trajectories $x^* = (S^*, J^*) : [t_0, t_1] \to X^n$ and $P^* : [t_0, t_1] \to \mathcal{L}(X^n, (X^k, \preccurlyeq))$ satisfying

$$\dot{x}^*(t) = f[x^*(t), u^*(t), t] \quad \text{a.e. } t \in [t_0, t_1] \tag{8}$$

$$x^*(t_0) = x_0 \tag{9}$$

$$S^*(t_1) = S_1 \tag{10}$$

$$\dot{P}^*(t) = -\frac{\partial H}{\partial x}[x^*(t), P^*(t), u^*(t), t] = -P^*(t)\frac{\partial f}{\partial x}[x^*(t), u^*(t), t] \quad \text{a.e. } t \in [t_0, t_1] \tag{11}$$

$$P^*_S(t_1) = P_S \tag{12}$$

$$P^*_J(t_1) = P_0 \tag{13}$$

the following condition holds

$$H[x^*(t), P^*(t), u^*(t), t] \preccurlyeq H[x^*(t), P^*(t), u(t), t]$$
$$\text{for all } u(t) \in \Omega \text{ and a.e. } t \in [t_0, t_1] \;, \tag{14}$$

i.e., the Hamiltonian is globally infimized with respect to $u(t)$ along the trajectory (x^*, P^*) defined by (8) through (13).

Remark II.6. Unless the superior control is somewhat singular, the operator P_0 can be chosen to be the identity operator, $P_0 = I \in \mathcal{L}((X^k, \preccurlyeq), (X^k, \preccurlyeq), \preccurlyeq)$, with no loss of generality. For further discussions, see Section II.4.

Theorem II.7. Solution to Infimization Problem II.2

In order that the admissible control u^* be a superior solution to the Infimization Problem II.2, it is necessary that there exist a positive map $P_0 \succ 0 \in \mathcal{L}((X^k,\preccurlyeq),(X^k,\preccurlyeq),\preccurlyeq)$ such that along the trajectories $x^* = (S^*,J^*)$ and P^* satisfying

$$\dot{x}^*(t) = f[x^*(t),u^*(t),t] \quad \text{a.e. } t \in [t_0,t_1] \tag{15}$$

$$x^*(t_0) = x_0 \tag{16}$$

$$\dot{P}^*(t) = -\frac{\partial H}{\partial x}[x^*(t),P^*(t),u^*(t),t] = -P^*(t)\frac{\partial f}{\partial x}[x^*(t),u^*(t),t] \quad \text{a.e. } t \in [t_0,t_1] \tag{17}$$

$$P^*_S(t_1) = 0 \tag{18}$$

$$P^*_J(t_1) = P_0 \tag{19}$$

the Hamiltonian is globally infimized, i.e.,

$$H[x^*(t),P^*(t),u^*(t),t] \preccurlyeq H[x^*(t),P^*(t),u(t),t]$$
$$\text{for all } u(t) \in \Omega \text{ and a.e. } t \in [t_0,t_1] \ . \tag{20}$$

Remark II.8. Clearly, the Remark II.6 applies to Theorem II.7 as well as to Theorem II.5. For further discussion, see Section II.4.

II.3. Proof of the Infimum Principle

The proof of the infimum principle stated in Theorems II.5 and II.7 closely parallels the proof of the maximum principle for scalar-valued cost functionals by Halkin [11]. The results which are independent of the partial ordering of the cost subspace of the state space are taken over without proof.

The major difficulty in the proof of the infimum principle stems from the fact that in the real line the complement of the positive cone is a convex set, whereas in an

integrally closed directed linear poset, the complement of the positive cone is not convex.

The choice of notation in the following proof should allow easy referencing to [11].

For an admissible control u, which fact is written $u \in F^*$, the solution of the differential equation (1), with initial condition (2), at time t is denoted by $x(t;u)$ and the corresponding whole trajectory x for t over $[t_0,t_1]$ by $x(u)$. For any admissible control u, the Fréchet derivative $D(t;u)$ is defined for all $t \in [t_0,t_1]$ by

$$D(t;u) = \left.\frac{\partial f(x,u(t),t)}{\partial x}\right|_{x = x(t;u)} . \tag{21}$$

By assumption, $D(t;u)$ is bounded and measurable over $[t_0,t_1]$. Furthermore, the transition operator $G(t;u)$ associated with $D(t;u)$ is introduced by the operator differential equation

$$\dot{G}(t;u) = - G(t;u)D(t;u) \quad \text{a.e. } t \in [t_0,t_1] \tag{22}$$

with boundary condition at t_1

$$G(t_1;u) = I, \tag{23}$$

where $I : X^n \to X^n$ is the identity operator. The transition operator $G(t;u)$ exists, is unique, bounded over $[t_0,t_1]$, and invertible for all $t \in [t_0,t_1]$.

For every trajectory $x(t;u^*)$ corresponding to a $u^* \in F^*$, a comoving space $Y^n(u^*)$ with elements y is defined by

$$y = G(t;u^*)(x-x(t;u^*)) , \quad x \in X^n. \tag{24}$$

In the comoving space $Y^n(u^*)$ of a trajectory $x(u^*)$; $u^* \in F^*$, a trajectory $x(u)$ corresponding to $u \in F^*$ is denoted by

$$y(u,u^*) = \{(y(t;u,u^*),t) : t \in [t_0,t_1]\} . \tag{25}$$

Furthermore, a "first-order approximation"

$$y^+(u,u^*) = \{(y^+(t;u,u^*),t) : t \in [t_0,t_1]\} \tag{26}$$

of $y(u,u^*)$ is defined for all $t \in [t_0,t_1]$ by

$$y^+(t;u,u^*) = \int_{t_0}^{t} G(s;u^*)\left[f(x(s;u^*),u(s),s) - f(x(s;u^*),u^*(s),s)\right]ds \quad . \tag{27}$$

$y^+(t;u,u^*)$ is a first-order approximation in the sense that the sup-norm of $y^+(u,u^*) - y(u,u^*)$ is bounded by a continuous, nondecreasing function $O(r)$, such that $\lim_{r \to 0} \frac{O(r)}{r} = 0$, where $r = \mu(\{t : u(t) \neq u^*(t),\ t \in [t_0,t_1]\})$ [11, p. 54] (μ= Lebesgue measure). For every pair of admissible controls u and u*, both $y(u,u^*)$ and $y^+(u,u^*)$ exist, are unique, and continuous.

Now the following "reachable sets" are introduced :

$$H = \{x(t_1;u) : u \in F^*\} \tag{28}$$

is the set of all final states reachable from the initial state (2) by applying an appropriate admissible control.

$$H(u^*) = \{y(t_1;u,u^*) : u \in F^*\} \text{ for any } u^* \in F^* \tag{29}$$

is the set H of (28) described in the coordinates of the comoving space $Y^n(u^*)$ of $u^* \in F^*$.

$$H^+(u^*) = \{y^+(t_1;u,u^*) : u \in F^*\} \text{ for any } u^* \in F^* \tag{30}$$

is the first-order approximation of $H(u^*)$ at $y(t_1;u^*,u^*) = 0 \in Y^n(u^*)$, and is known to be convex for every $u^* \in F^*$ [11, p. 70]. Observe, that by (23) and (24), the map from H to $H(u^*)$ is simply a translation, which depends on u*, i.e.,

$$H(u^*) = \{x - x(t_1;u^*) : x \in H\} \ . \tag{31}$$

<u>Proposition II.9.</u> *If the admissible control u* is a superior solution to Infimization Problem II.1 or II.2, then the point $x(t_1;u^*)$ is a boundary point of the set H in X^n.*

<u>Proof:</u> If $x(t_1;u^*) = (S(t_1;u^*),J(t_1;u^*))$ is in the interior of H, then there exists $\varepsilon > 0$, sucht that $z = (S(t_1;u^*),J(t_1;u^*) - J_\varepsilon)$ is also in H, where $(0,J_\varepsilon) \in X^n$ has norm ε and $J_\varepsilon \succ 0 \in (X^k,\preccurlyeq)$ is a positive element of the cost subspace. Hence, there exists a $u_z \in F^*$ satisfying the boundary constraints and resulting in the cost $J(t_1;u_z) \prec J(t_1;u^*)$.

Thus, $x(t_1;u^*)$ has to lie on the boundary of H, in order that u^* be superior (or noninferior, for that matter [12]).

Remark II.10. Clearly, by (31), every boundary point of H is also a boundary point of $H(u^*)$ for any $u^* \in F^*$. In particular, if $x(t_1;u^*)$ is a boundary point of H, then $y(t_1;u^*,u^*) = 0$ is a boundary point of the set $H(u^*)$.

Proposition II.11. If the point $y = 0$ is a boundary point of the set $H(u^*)$, then the point $y = 0$ is also a boundary point of the set $H^+(u^*)$.

Proof: See [11, p. 77].

Proposition II.12. If the point $y = 0$ is a boundary point of the set $H^+(u^*)$, then there exists a nonzero linear map $P^*(t_1;u^*) : Y^n(u^*) \to (X^k, \preccurlyeq)$, such that

$$P^*(t_1;u^*)G(t;u^*)[f(x(t;u^*),u(t),t) - f(x(t;u^*),u^*(t),t)] \succcurlyeq 0$$

$$\text{for all } u(t) \in \Omega \text{ and a.e. } t \in [t_0,t_1]. \tag{32}$$

Proof: Since $H^+(u^*)$ is convex and $y = 0$ is a boundary point of $H^+(u^*)$, there exists a nonzero linear functional $p(t_1;u^*) : Y^n(u^*) \to (R,\leqslant)$, such that $p(t_1;u^*)y \geqslant 0$ for all $y \in H^+(u^*)$. Consequently, there exists a nonzero linear map $P^*(t_1;u^*) : Y^n(u^*) \to (X^k, \preccurlyeq)$, such that

$$P^*(t_1;u^*)y \succcurlyeq 0 \text{ for all } y \in H^+(u^*) \; . \tag{33}$$

Assume, that there exists a control u satisfying the constraint $u(t) \in \Omega$ for all $t \in [t_0,t_1]$, such that

$$J_1(t) = P^*(t_1;u^*)G(t;u^*)[f(x(t;u^*),u(t),t) - f(x(t;u^*),u^*(t),t)] \not\succcurlyeq 0 \tag{34}$$

for $t \in E$, where E is a Borel set in $[t_0,t_1]$ of positive measure, $\mu(E) > 0$. In order to conclude from this assumption, that there exists $y \in H^+(u^*)$ not satisfying (33), a Borel subset $F \subseteq E$

is constructed, such that $J_1(t)$ of (34) belongs to one and the same open half-space of $(X^k,\preccurlyeq)$ for all $t \in F$, which is separated from the positive cone of $(X^k,\preccurlyeq)$:

Since the complement of the positive cone of $(X^k,\preccurlyeq)$, $\{J \in (X^k,\preccurlyeq) : J \not\succcurlyeq 0\}$, is an open set, there exists $\varepsilon > 0$, such that the angular distance d between $J_1(t)$ of (34) and the positive cone is bigger than ε, i.e.,

$$\inf_{\substack{t \in E \\ J_2 \succcurlyeq 0}} d(J_1(t), J_2) \overset{\Delta}{=} \inf_{t \in E} \min_{J_2 \succcurlyeq 0} \frac{\| J_1(t) - J_2 \|}{\| J_1(t) \|} > \varepsilon , \tag{35}$$

where $\|.\|$ denotes the norm of $(X^k,\preccurlyeq)$. Also, since $(X^k,\preccurlyeq)$ is finite-dimensional, there exists a closed polyhedral cone Q in $(X^k,\preccurlyeq)$, which is circumscribed to the positive cone, such that the angular distance d between any element of Q and the positive cone is at most ε, i.e.,

$$\sup_{\substack{J_1 \in Q \\ J_1 \neq 0}} \min_{J_2 \succcurlyeq 0} d(J_1, J_2) = \sup_{\substack{J_1 \in Q \\ J_1 \neq 0}} \min_{J_2 \succcurlyeq 0} \frac{\| J_1 - J_2 \|}{\| J_1 \|} \leq \varepsilon . \tag{36}$$

Furthermore, since $\mu(E) > 0$ and Q has finitely many faces, there exists a hyperplane $\Pi(Q)$, viz. a face of Q, such that $J_1(t)$ of (34) lies in the open half space $R(\Pi(Q))$, defined by $\Pi(Q)$ and not containing Q, for all $t \in F$, where F is a Borel subset of E of positive measure, $\mu(F) > 0$.

Since $R(\Pi(Q))$ is an open half space of $(X^k,\preccurlyeq)$ containing $J_1(t)$ for all $t \in F$, there exists $\varepsilon_1 > 0$, such that the distance of $J_1(t)$ from the hyperplane $\Pi(Q)$ is always greater than ε_1, i.e.,

$$\inf_{\substack{t \in F \\ J_2 \in \Pi(Q)}} \| J_1(t) - J_2 \| > \varepsilon_1 \quad . \tag{37}$$

Denoting by $\chi(F)$ the support function of F, it follows from the definition of $\Pi(Q)$ and (37)

that the element

$$\Delta J = \int_{t \in F} P^*(t_1;u^*)G(t;u^*)[f(x(t;u^*),u(t),t) - f(x(t;u^*),u^*(t),t)]dt \tag{38}$$

lies in the complement of the positive cone of $(X^k,\preccurlyeq)$ together with a ball of radius $\varepsilon_1\mu(F)$ and center ΔJ, hence that the admissible control [11, p. 45]

$$v(t) = u^*(t) + \chi(t)(u(t) - u^*(t)) \tag{39}$$

generates $y^+(t_1;v,u^*) \in H^+(u^*)$, such that

$$P^*(t_1;u^*)y^+(t_1;v,u^*) = \Delta J \not\succcurlyeq 0 \quad , \tag{40}$$

contradicting (33).

This concludes the proof of Proposition II.12.

Since X^n and $Y^n(u^*)$ are isomorphic, viz. by the translation relationship (31), $P^*(t_1;u^*)$ can be split up into $P^*(t_1;u^*) = (P^*_S(t_1;u^*),P^*_J(t_1;u^*))$, where $P^*_J(t_1;u^*) \in \mathcal{L}((X^k,\preccurlyeq),(X^k,\preccurlyeq),\preccurlyeq)$ and $P^*_S(t_1;u^*) \in \mathcal{L}((X^k,\preccurlyeq)^\perp,(X^k,\preccurlyeq))$.

Proposition II.13. In Proposition II.12, the linear map $P^*(t_1;u^*)$ is such that $P^*_J(t_1;u^*)$ is a positive map, i.e.,

$$P^*_J(t_1;u^*) \succcurlyeq 0 \in \mathcal{L}((X^k,\preccurlyeq),(X^k,\preccurlyeq),\preccurlyeq) \quad . \tag{41}$$

Proof: In the case of the free-end-point Infimization Problem II.2, the set $H^+(u^*)$ must not intersect the open subset

$$V = \{y = (S,J) \in Y^n(u^*) : J \not\succcurlyeq 0,\ S \in (X^k,\preccurlyeq)^\perp\} \tag{42}$$

of $Y^n(u^*)$, because otherwise a $v \in F^*$ could be found, which would be noninferior to u^*, since $H^+(u^*)$ is the first order convex approximation of $H(u^*)$ at $y(t_1;u^*,u^*) = 0 \in Y^n(u^*)$. Hence,

$$H^+(u^*) \subseteq \{y = (S,J) \in Y^n(u^*) : J \succcurlyeq 0,\ S \in (X^k,\preccurlyeq)^\perp\} \tag{43}$$

and therefore every linear map $P \in \mathcal{L}(Y^n(u^*),(X^k,\preccurlyeq))$ of the form $P = (P_S,P_J)$ with

$$P_S = 0 \in \mathcal{L}((X^k,\preccurlyeq)^{\perp},(X^k,\preccurlyeq)) \tag{44}$$

and $P_J \succcurlyeq 0 \in \mathcal{L}((X^k,\preccurlyeq),(X^k,\preccurlyeq),\preccurlyeq)$ satisfies

$$Py \succcurlyeq 0 \text{ for all } y \in H^+(u^*)\,. \tag{45}$$

In the case of the fixed-end-point Infimization Problem II.1, the interior of the set $H^+(u^*)$ must not intersect the subset

$$V = \{y = (S,J) \in Y^n(u^*) : J \not\succcurlyeq 0,\ S = 0\} \tag{46}$$

of $Y^n(u^*)$, because otherwise a $v \in F^*$ could be found, which would be noninferior to u^* since $H^+(u^*)$ is the first order convex approximation of $H(u^*)$ at $y(t_1;u^*,u^*) = 0 \in Y^n(u^*)$. Hence, $H^+(u^*)$ is contained in a set A of the form of the vector sum

$$A = B \oplus C \quad , \tag{47}$$

where $B = \{y = (S,J) \in Y^n(u^*) : S = 0,\ J \succcurlyeq 0\}$ and C is an n-k dimensional subspace of $Y^n(u^*)$ which is separated from B.

If B and C are disjoint, except at the origin, rather than merely separated, the Infimization Problem II.1 is called regular or nonsingular. In this case, every linear map $P \in \mathcal{L}(Y^n(u^*),(X^k,\preccurlyeq))$ of the form $P = (P_S,P_J)$ with $\mathcal{N}(P) = C$ and $P_J = I$ satisfies (45) for all $y \in H^+(u^*)$.

If $B \cap C$ contains more than the origin, i.e., if B and C are separated but not dis= joint, the Infimization Problem II.1 is called singular. In this case, there exists a linear map $P_0 \succcurlyeq 0 \in \mathcal{L}((X^k,\preccurlyeq),(X^k,\preccurlyeq),\preccurlyeq)$ (but $P_0 \neq I$), such that every linear map $P \in \mathcal{L}(Y^n(u^*),(X^k,\preccurlyeq))$ of the form $P = (P_S,P_J)$ with $\mathcal{N}(P) \supset C$ and $P_J = P_0$ satisfies (45) for all $y \in H^+(u^*)$.

This concludes the proof of Proposition II.13.

Remark II.14. If the linear map $P^*(t;u^*) \in \mathcal{L}(X^n,(X^k,\preccurlyeq))$ is defined by

$$P^*(t;u^*) = P^*(t_1;u^*)G(t;u^*) \tag{48}$$

then by (22) and (23), $P^*(t;u^*)$ satisfies the operator differential equation

$$\dot{P}^*(t;u^*) = -P^*(t;u^*)D(t;u^*) \quad \text{a.e. } t \in [t_0,t_1] \tag{49}$$

with boundary condition

$$P^*(t;u^*) = P^*(t_1;u^*) \text{ for } t = t_1 \;. \tag{50}$$

Furthermore, with the definition of the Hamiltonian (7), the equation (32) of Proposition II.12 can be written as

$$H[x^*(t),P^*(t;u^*),u^*(t),t] \preccurlyeq H[x^*(t),P^*(t;u^*),u(t),t]$$
$$\text{for all } u(t) \in \Omega \text{ and a.e. } t \in [t_0,t_1] \;. \tag{51}$$

Remark II.15. Combining Proposition II.9, Remark II.10, Propositions II.11, II.12, and II.13, equation (44), Remark II.14, and equation (21) completes the proofs of the Theorems II.5 and II.7.

II.4. Discussion

The infimum principle in the two Theorems II.5 and II.7 is stated for a globally superior control u^* as required in the Infimization Problems II.1 and II.2, respectively. In the proof of the infimum principle in Section II.3, the analysis is global with respect to the control $u(t) \in \Omega$ as expressed in (32), (14), and (20). However, the analysis is only local in the state space X^n and in the space of comoving coordinates $Y^n(u^*)$, since the first order convex approximation $H^+(u^*)$ of the reachable set $H(u^*)$ is investigated ((26) through (30)).

Therefore, the infimum principle also applies to a control u^*, which is locally superior in the following sense:

Definition II.16. An admissible control u^* is locally superior, if there exists $\varepsilon > 0$, such that for every admissible control u satisfying the boundary constraint at t_1 (i.e., (4) in Problem II.1) and generating $x(t_1;u)$ in the ε-neighbourhood of $x(t_1;u^*)$ (i.e., $\|x(t_1;u) - x(t_1;u^*)\| \leq \varepsilon$), the cost component $J(t_1;u)$ of $x(t_1;u)$ is related to the locally superior cost $J(t_1;u^*)$ by $J(t_1;u^*) \preccurlyeq J(t_1;u)$.

Clearly, a globally superior control is also locally superior but not vice versa. Therefore, when applying the infimum principle to solving infimization problems, the globality of the superior solution has to be verified separately, in addition to investigating the existence of a superior solution.

In Theorems II.5 and II.7, the costate $P(t)$ was chosen to be a linear map from X^n into the cost-subspace $(X^k,\preccurlyeq)$ of X^n, i.e., $P(t) \in \mathcal{L}(X^n,(X^k,\preccurlyeq))$. Now, for j, $1 \leq j < k$, let $(X^j,\preccurlyeq)$ denote a j-dimensional integrally closed directed linear poset, in particular for $j = 1$, $(X^1,\preccurlyeq) = (R,\leq)$. If u^*, x^*, and P^* satisfy the infimum principle, then for every positive map $P_1 \in \mathcal{L}((X^k,\preccurlyeq),(X^j,\preccurlyeq),\preccurlyeq)$, $P_1 \succcurlyeq 0$,

$$\begin{aligned} &P_1 H[x^*(t),P^*(t),u^*(t),t] = P_1 P^*(t) f[x^*(t),u^*(t),t] \\ &\preccurlyeq P_1 H[x^*(t),P^*(t),u(t),t] = P_1 P^*(t) f[x^*(t),u(t),t] \\ &\text{for all } u(t) \in \Omega \text{ and a.e. } t \in [t_0,t_1] \quad . \end{aligned} \tag{52}$$

By the linearity of the costate differential equation (11) or (17), respectively, this implies that the costate map $P(t)$ could have been chosen in $\mathcal{L}(X^n,(X^j,\preccurlyeq))$ for any $1 \leq j < k$ rather than in $\mathcal{L}(X^n,(X^k,\preccurlyeq))$. However, the infimum principle for $j < k$ would be weaker than that of Theorems II.5 and II.7, because the truth of (52) for a particular positive map P_1 does not, in general, imply the truth of (52) for all positive maps P_1 (see Lemma A.2). As a matter of fact, the conditions obtained for $j = 1$ merely constitute necessary conditions for noninferiority [6], [7], [8], [14].

In the proof of Proposition II.13 and in the Remark II.6, an infimization problem has been called regular or nonsingular, if P_0 in (13) or (19), respectively, can be taken to be the identity operator, otherwise the infimization problem has been called singular. In optimal control theory for scalar-valued cost functionals, this type of singularity of an optimal control problem is well known [1], [13]. In this case, the nonnegative constant p_0 happens to be zero. It can be expected that singularity of superior controls does or does not occur under conditions quite analogous to the conditions, under which singularity of optimal controls for scalar-valued cost functionals does or does not occur, respectively.

In the statements of the Infimization Problems II.1 and II.2 in Section II.1, the space $(X^k,\preccurlyeq)$ has been chosen to be a subspace of the state space X^n, and the dynamics (1) of the system have been allowed to depend on all components of the state $x(t) = (S(t),J(t))$ (3). This is the most general case.

In the important special case, where $f[x,u,t]$ in (1) actually does not depend on the cost component $J(t)$ of the state $x(t)$, the component $P_J(t)$ of the costate $P(t) = (P_S(t),P_J(t))$ (Convention II.3) is constant, hence, $P_J(t) \equiv P_0$ in Theorems II.5 and II.7. It then is more convenient to use the following convention :

<u>Convention II.17.</u> The cost space is a k-dimensional integrally closed directed linear poset $(X^k,\preccurlyeq)$, the state space is an n-dimensional linear space X^n (previously denoted by $(X^k,\preccurlyeq)^\perp$), of which $(X^k,\preccurlyeq)$ is not a subspace. The state is $x(t) \in X^n$ (previously $S(t)$). The costate $P(t)$ is a linear map from the state space into the cost space, i.e., $P(t) \in \mathcal{L}(X^n,(X^k,\preccurlyeq))$. In addition, there now is a constant linear map $P_0 \in \mathcal{L}((X^k,\preccurlyeq),(X^k,\preccurlyeq),\preccurlyeq)$ (previously $P_J(t_1)$ in Theorems II.5 and II.7).

In this special case and using Convention II.17, the Infimization Problem II.1 is

restated as follows :

Infimization Problem II.18. Given the system

$$\dot{x}(t) = f_1[x(t),u(t),t] \quad \text{a.e. } t \in [t_0,t_1] \tag{53}$$

$$x(t_0) = x_0 \tag{54}$$

$$x(t_1) = x_1 \quad , \tag{55}$$

find an admissible control u^* (in particular, $u^*(t) \in \Omega$ for all $t \in [t_0,t_1]$), such that the cost $J(u)$ defined by

$$J(u) = \int_{t_0}^{t_1} f_2[x(t),u(t),t]dt \tag{56}$$

is globally infimized, i.e.,

$$J(u^*) \preccurlyeq J(u) \tag{57}$$

for all admissible controls u meeting the boundary constraint $x(t_1) = x_1$.

Of course, the statement of the Infimization Problem II.2 can be adapted in a similar way.

With the new Hamiltonian

$$H[x(t),P_0,P(t),u(t),t] = P_0 f_2[x(t),u(t),t] + P(t)f_1[x(t),u(t),t] \tag{58}$$

instead of (7), Theorem II.5 applied to the Infimization Problem II.18 becomes

Corrolary II.19. In order that the admissible control u^* be a superior solution to the Infimization Problem II.18, it is necessary that there exist a nonzero map $(P^*(t_1),P_0^*)$ $\in \mathcal{L}(X^n \times (X^k,\preccurlyeq),(X^k,\preccurlyeq))$ with P_0^* positive, such that along the trajectories x^*, P^* satisfying

$$\dot{x}^*(t) = f_1[x^*(t),u^*(t),t] \quad \text{a.e. } t \in [t_0,t_1] \tag{59}$$

$$x^*(t_0) = x_0 \tag{60}$$

$$x^*(t_1) = x_1 \tag{61}$$

$$\dot{P}^*(t) = -\frac{\partial H}{\partial x}[x^*(t), P_0^*, P^*(t), u^*(t), t]$$

$$= -P_0^* \frac{\partial f_2}{\partial x}[x^*(t), u^*(t), t] - P^*(t)\frac{\partial f_1}{\partial x}[x^*(t), u^*(t), t] \quad \text{a.e. } t \in [t_0, t_1] \qquad (62)$$

the following condition holds

$$H[x^*(t), P_0^*, P^*(t), u^*(t), t] \preccurlyeq H[x^*(t), P_0^*, P^*(t), u(t), t]$$

$$\text{for all } u(t) \in \Omega \text{ and a.e. } t \in [t_0, t_1] \quad , \qquad (63)$$

i.e., the Hamiltonian is globally infimized with respect to $u(t)$ along the trajectory (x^*, P^*) defined by (59) through (63).

Of course, the Corollary II.19 specializes to exactly the type of theorems given in [1] for the case $k = 1$ (scalar-valued cost). In the nonsingular case, again, $P_0^* = I \in \mathcal{L}((X^k, \preccurlyeq), (X^k, \preccurlyeq), \preccurlyeq)$.

The reader should have no trouble in extending the necessary conditions for a control to be optimal with respect to a scalar-valued cost [1] in problems with other boundary conditions or other types of cost functionals than those reported here, to the corresponding set of necessary conditions for a control to be superior with respect to a non-scalar-valued cost (transversality conditions, final state penalty terms, free final time, etc.).

III. SUFFICIENCY RESULTS

The purpose of this section is to show how some of the known results in the theory of optimal control for scalar-valued cost functionals concerning sufficiency conditions for a control to be optimal [13] can be extended to the theory of superior controls for non-scalar-valued performance criteria.

In Section III.1, a Hamilton-Jacobi-Bellman type of theory is developed, which constitutes a sufficiency condition for a control to be superior relative to a region Z in the product of the state space X^n and the time axis R. This theory is quite analogous to the Hamilton-Jacobi-Bellman theory in optimal control with scalar-valued cost.

In Section III.2, some sufficiency results for global optimality are generalized to sufficiency results for global superiority.

On the other side, all of the non-trivial existence results of optimal control for scalar-valued cost (e.g. [13, pp. 259 ff. and 286 ff.] and [5]) generalize to existence results for the less restrictive noninferior controls (see [15]) rather than for the more restrictive superior control, in the case of non-scalar-valued performance criteria.

III. 1. Hamilton-Jacobi-Bellman Theory

Consider the dynamic system

$$\dot{x}(t) = f[x(t), u(t), t] \quad , \tag{1}$$

where the state $x(t) \in X^n$, the control $u(t)$ is restricted to a closed subset $\Omega \subset X^m$, and $f : X^n \times \Omega \times R \to X^n$ is continuously differentiable on its domain. A control u is admissi= ble, if it is piecewise continuous and satisfies $u(t) \in \Omega$ for all t of interest.

It is assumed, that for every admissible control u and any initial state

$$x(t_0) = x_0 \quad , \tag{2}$$

there exists an absolutely continuous solution x satisfying (1) a.e. and (2), (which then is automatically unique by the local Lipschitz continuity of f with respect to x(t)).

The target set S is a subset of $X^n \times \{t \in R \mid t > t_0\}$. The non-scalar-valued cost to be infimized is of the form

$$J(u) = K[x(t_1),t_1] + \int_{t_0}^{t_1} L[x(t),u(t),t]dt \tag{3}$$

and attains its value in the integrally closed directed k-dimensional linear poset $(X^k,\preccurlyeq)$, which may or may not be a subspace of the state space X^n. Here, L and K are continuously differentiable on their respective domains, and the final time $t_1 > t_0$ is the first time, the target set S is met by the trajectory x generated by the admissible control u.

Convention III.1. Z denotes a connected subset of $X^n \times R$, which intersects the target set S. Furthermore, Π always denotes a connected subset of $\mathcal{L}(X^n,(X^k,\preccurlyeq)) \times R$.

Definition III.2. The Hamiltonian $H : X^n \times \mathcal{L}(X^n,(X^k,\preccurlyeq)) \times \Omega \times R \rightarrow (X^k,\preccurlyeq)$ defined by

$$H(x,P,u,t) = L(x,u,t) + Pf(x,u,t) \tag{4}$$

is called normal relative to Z and Π, if for each (x,P,t) with $(x,t) \in Z$ and $(P,t) \in \Pi$, the Hamiltonian has a unique absolute infimum with respect to all $u \in \Omega$, viz. at the point

$$u = u^0(x,P,t) \in \Omega \quad , \tag{5}$$

which is called the H-infimal control relative to Z and Π.

Now, the following theorem can be stated, which is the analogue to the Hamilton-Jacobi-Bellman theory of optimal control for scalar-valued cost functionals [1, p. 351], [3, p. 315].

<u>Theorem III.3</u>. Suppose that the Hamiltonian (4) is normal relative to regions Z and $\overline{\Pi}$ and let $u^0(x,P,t)$ denote the H-infimal control relative to Z and $\overline{\Pi}$. Suppose that u^* is an admissible control, which transfers $(x_0,t_0) \in Z$ to S, such that the corresponding trajectory x^* stays entirely in Z, i.e.,

$$(x^*(t),t) \in Z \text{ for all } t \in [t_0,t_1] \quad . \tag{6}$$

Suppose that there exists a continuously differentiable function $J(x,t)$ on Z satisfying the partial differential equation

$$\frac{\partial J}{\partial t}(x,t) + H\left[x,\frac{\partial J}{\partial x}(x,t),u^0\left[x,\frac{\partial J}{\partial x}(x,t),t\right],t\right] = 0 \tag{7}$$

with boundary condition

$$J(x,t) = K(x,t) \text{ for } (x,t) \in S \cap Z \quad , \tag{8}$$

such that $\frac{\partial J}{\partial x}(x,t)$ stays entirely in $\overline{\Pi}$, i.e.,

$$\left(\frac{\partial J}{\partial x}(x,t),t\right) \in \overline{\Pi} \text{ for all } t \in [t_0,t_1] \quad . \tag{9}$$

If

$$u^*(t) = u^0\left[x^*(t),\frac{\partial J}{\partial x}(x^*(t),t),t\right] \text{ for all } t \in [t_0,t_1] \quad , \tag{10}$$

then $u^* : [t_0,t_1] \to \Omega$ is a superior control relative to the set $U_{Z\overline{\Pi}}$ of admissible controls u generating trajectories x and $\frac{\partial J}{\partial x}$ lying entirely in Z and $\overline{\Pi}$, respectively, and the infimal cost is given by

$$J(u^*) = J(x_0,t_0) \preccurlyeq J(u) \quad \text{for all } u \in U_{Z\overline{\Pi}} \quad . \tag{11}$$

<u>Remark III.4</u>. Clearly, if the Theorem III.3 applies to the case of the initial point (x_0,t_0) $\in Z$, then, it also applies to any initial point $(x^*(t),t)$ with $t \in [t_0,t_1]$ and the corresponding "cost to go" $J(x^*(t),t)$.

Corollary III.5. If $Z = X^n \times [t_a, t_b]$, and $S \subset Z$, and $\Pi = \mathcal{L}(X^n, (X^k, \preccurlyeq)) \times [t_a, t_b]$, then the Theorem III.3 provides a sufficiency condition for global superiority for all initial states $(x_0, t_0) \in Z$.

Corollary III.6. If in addition to the assumptions of Theorem III.3, the function $J(x,t)$ is twice continuously differentiable on Z, Z is open in $X^n \times R$, and the H-infimal control $u^0[x, \frac{\partial J}{\partial x}(x,t), t]$ is continuously differentiable with respect to x and continuous in t, then the function $P : [t_0, t_1] \to \mathcal{L}(X^n, (X^k, \preccurlyeq))$ defined by $P(t) = \frac{\partial J}{\partial x}(x^*(t), t)$ is a costate function in the infimization problem (1), (2), (3). This follows directly from the corresponding proof by Kalman in [3, p. 320] at least in the case where $\Omega = X^m$.

Proof of Theorem III.3. With $x = x^*(t)$, (7) becomes

$$
\begin{aligned}
&\frac{\partial J}{\partial t}(x^*(t),t) + H\left[x^*(t), \frac{\partial J}{\partial x}(x^*(t),t), u^0\left[x^*(t), \frac{\partial J}{\partial x}(x^*(t),t),t\right],t\right] \\
&= \frac{\partial J}{\partial t}(x^*(t),t) + L\left[x^*(t), u^0\left[x^*(t), \frac{\partial J}{\partial x}(x^*(t),t),t\right],t\right] \\
&\quad + \frac{\partial J}{\partial x}(x^*(t),t) f\left[x^*(t), u^0\left[x^*(t), \frac{\partial J}{\partial x}(x^*(t),t),t\right],t\right] \\
&= \frac{dJ}{dt}(x^*(t),t) + L\left[x^*(t), u^0\left[x^*(t), \frac{\partial J}{\partial x}(x^*(t),t),t\right],t\right] = 0 \quad . \qquad (12)
\end{aligned}
$$

Integrating (12) from t_0 to t_1 and using (8) yields

$$
\begin{aligned}
&J(x^*(t_1),t_1) - J(x_0,t_0) + \int_{t_0}^{t_1} L\left[x^*(t), u^0\left[x^*(t), \frac{\partial J}{\partial x}(x^*(t),t),t\right],t\right]dt \\
&= K(x^*(t_1),t_1) - J(x_0,t_0) + \int_{t_0}^{t_1} L\left[x^*(t), u^0\left[x^*(t), \frac{\partial J}{\partial x}(x^*(t),t),t\right],t\right]dt \\
&= J(u^*) - J(x_0,t_0) = 0 \quad . \qquad (13)
\end{aligned}
$$

For any admissible control $u : [t_0, t_2] \to \Omega$ transferring (x_0, t_0) to $(x_u(t_2), t_2) \in S$, such that

the trajectory $x_u : [t_0, t_2] \to X^n$ lies entirely in Z and $\frac{\partial J}{\partial x}(x_u(.),.) : [t_0, t_2] \to \mathcal{L}(X^n, (X^k, \preccurlyeq))$ lies entirely in Π, equation (7) reads at every $t \in [t_0, t_2]$ and with $x = x_u(t)$,

$$\frac{\partial J}{\partial t}(x_u(t),t) + H\left[x_u(t), \frac{\partial J}{\partial x}(x_u(t),t), u^0\left[x_u(t), \frac{\partial J}{\partial x}(x_u(t),t),t\right],t\right] = 0 \quad . \tag{14}$$

And by the normality of H, (14) yields

$$\begin{aligned} &\frac{\partial J}{\partial t}(x_u(t),t) + H\left[x_u(t), \frac{\partial J}{\partial x}(x_u(t),t), u(t), t\right] \\ &= \frac{\partial J}{\partial t}(x_u(t),t) + L[x_u(t),u(t),t] + \frac{\partial J}{\partial x}(x_u(t),t) f[x_u(t),u(t),t] \\ &\begin{cases} \succ 0\,, \text{ whenever } u(t) \neq u^0[x_u(t), \frac{\partial J}{\partial x}(x_u(t),t),t] \\ = 0\,, \text{ whenever } u(t) = u^0[x_u(t), \frac{\partial J}{\partial x}(x_u(t),t),t] \end{cases} \quad . \end{aligned} \tag{15}$$

Integrating (15) from t_0 to t_2 and using (8) then results in

$$J(u) - J(x_0, t_0) \succcurlyeq 0 \quad . \tag{16}$$

Subtracting (13) from (16) results in

$$J(u) - J(u^*) \succcurlyeq 0 \quad \text{for all } u \in U_{Z\Pi} . \tag{17}$$

Combining (13) and (17) completes the proof of (11) and of the Theorem III.3.

III.2. More Sufficiency Results

In this section a sufficiency result given by Lee and Markus [13, p. 341] is generalized to the case of non-scalar-valued performance criteria.

Consider the dynamic system

$$\dot{x}(t) = A(t)x(t) + h(u(t),t) \quad , \quad \text{a.e.} \quad , \tag{18}$$

where the state $x(t) \in X^n$, the control $u(t) \in \Omega \subset X^m$ is measurable, $A(t) \in \mathcal{L}(X^n, X^n)$ is continuous in t over $[t_0, t_1]$, and $h : \Omega \times [t_0, t_1] \to X^n$ is continuous on its domain. The initial state at the fixed initial time t_0 is

$$x(t_0) = x_0 \quad . \tag{19}$$

The target set $S \subset X^n$ at the fixed final time t_1 is closed and convex (and possibly $S = X^n$),

$$x(t_1) \in S \quad . \tag{20}$$

The non-scalar-valued cost to be infimized is of the form

$$J(u) = \int_{t_0}^{t_1} (f^0(x(t),t) + h^0(u(t),t))dt \tag{21}$$

and attains its value in the integrally closed directed linear poset $(X^k,\preccurlyeq)$. In (21), f^0, $\partial f^0/\partial x$, and h^0 are assumed to be continuous on their domains. Furthermore, f^0 is assumed to be convex in x (Definition II.19 of [2]) for all $t \in [t_0,t_1]$, i.e.,

$$f_0(sx_1 + (1-s)x_2,t) \preccurlyeq sf^0(x_1,t) + (1-s)f^0(x_2,t)$$
$$\text{for all } s \in [0,1], \text{ all } x_1 \text{ and } x_2 \text{ in } X^n, \text{ and all } t \in [t_0,t_1] \quad . \tag{22}$$

<u>Theorem III.7.</u> If the optimal control problem is nonsingular and if there exists a measurable control $u^*: [t_0,t_1] \to \Omega$ satisfying the necessary conditions of the infimum principle, then u^* is globally superior (although not necessarily unique), provided that one of the following conditions holds :

a) $S = X^n$ or $x^*(t_1) \in \text{int}(S)$

b) $x^*(t_1) \in \partial S$ and $P^*(t_1)(x - x^*(t_1)) \preccurlyeq 0$

for all $x \in S$, which are reachable at t_1

c) $S = \{x_1\}$. (23)

<u>Proof.</u> Since the optimal control problem is nonsingular, $P_0^* \in \mathcal{L}((X^k,\preccurlyeq),(X^k,\preccurlyeq),\preccurlyeq)$ can be taken to be the identity map, $P_0^* = I$, by Remark II.6. Hence, the Hamiltonian is

$$H(x,P,u,t) = f^0(x,t) + h^0(u,t) + PA(t)x + Ph(u,t) \quad . \tag{24}$$

Now, define the quantity $x^0(t) \in (X^k,\preccurlyeq)$ for all $t \in [t_0,t_1]$ by

$$\dot{x}^0(t) = f^0(x(t),t) + h^0(u(t),t) \quad \text{a.e. } t \in [t_0,t_1] \tag{25}$$

$$x^0(t_0) = 0 \ , \tag{26}$$

where $x : [t_0,t_1] \to X^n$ is the trajectory generated by the admissible control $u : [t_0,t_1] \to \Omega$, and consider the quantity

$$x^0(t) + P^*(t)x(t) \in (X^k, \preccurlyeq) \tag{27}$$

where $P^*(.)$ is the optimal costate, which satisfies

$$\dot{P}^*(t) = -\frac{\partial f^0}{\partial x}(x^*(t),t) - P^*(t)A(t) \quad \text{a.e. } t \in [t_0,t_1] \tag{28}$$

$$P^*(t_1) = P_1^* \quad , \tag{29}$$

according to the infimum principle.

Differentiating (27) and using (18) and (28)

$$\begin{aligned} &\frac{d}{dt}(x^0(t) + P^*(t)x(t)) = \dot{x}^0(t) + \dot{P}^*(t)x(t) + P^*(t)\dot{x}(t) \\ &= f^0(x(t),t) + h^0(u(t),t) - \frac{\partial f^0}{\partial x}(x^*(t),t)x(t) + P^*(t)h(u(t),t) \quad . \end{aligned} \tag{30}$$

Integrating (30) with respect to t over $[t_0,t_1]$ yields

$$\begin{aligned} &x^0(t_1) + P^*(t_1)x(t_1) - P^*(t_0)x_0 \\ &= \int_{t_0}^{t_1} \{f^0(x(t),t) + h^0(u(t),t) - \frac{\partial f^0}{\partial x}(x^*(t),t)x(t) + P^*(t)h(u(t),t)\}dt \quad . \end{aligned} \tag{31}$$

Evaluating (31) for u^* and its corresponding state x^* and subtracting this equation from (31),

$$\begin{aligned} &x^0(t_1) - x^{0*}(t_1) + P^*(t_1)(x(t_1) - x^*(t_1)) \\ &= \int_{t_0}^{t_1} \{ f^0(x(t),t) - f^0(x^*(t),t) - \frac{\partial f^0}{\partial x}(x^*(t),t)(x(t) - x^*(t)) \\ &\qquad + h^0(u(t),t) + P^*(t)h(u(t),t) - h^0(u^*(t),t) - P^*(t)h(u^*(t),t)\}dt \quad . \end{aligned} \tag{32}$$

The convexity and the differentiability of f^0 in x together with Lemma II.20 of [2] imply

$$f^0(x(t),t) - f^0(x^*(t),t) - \frac{\partial f^0}{\partial x}(x^*(t),t)(x(t) - x^*(t)) \succcurlyeq 0 \quad \text{for all } t \in [t_0,t_1] \ . \tag{33}$$

By hypothesis, u^* globally infimizes the Hamiltonian $H(x^*(t),P^*(t),u,t)$ for a.e. $t \in [t_0,t_1]$, which implies

$$h^0(u(t),t) + P^*(t)h(u(t),t) - h^0(u^*(t),t) - P^*(t)h(u^*(t),t) \succcurlyeq 0$$

$$\text{for a.e. } t \in [t_0,t_1] \quad . \tag{34}$$

Combining (32), (33), and (34) yields

$$x^0(t_1) - x^{0*}(t_1) + P^*(t_1)(x(t_1) - x^*(t_1)) \succcurlyeq 0 \quad . \tag{35}$$

In case (a), where $S = X^n$ or $x^*(t_1) \in \text{int } S$, necessarily $P^*(t_1) = 0$ in (29) and (35). Hence,

$$J(u) - J(u^*) = x^0(t_1) - x^{0*}(t_1) \succcurlyeq 0 \quad \text{for all admissible } u \text{ .} \tag{36}$$

In case (b), where $x^*(t_1) \in \partial S$, the hypothesis $P^*(t_1)(x - x^*(t_1)) \preccurlyeq 0$ for all $x \in S$, which are reachable at t_1 combines with (35) to (36). In case (c), where $S = \{x_1\}$, only the controls u meeting the boundary condition $x(t_1) = x_1 = x^*(t_1)$ are admissible. Thus, (35) reduces to (36) again.

If in the optimal control problem (18), . . . , (21), the target set $S = X^n$ and the non-scalar-valued cost to be infimized is of the form

$$J(u) = K(x(t_1)) + \int_{t_0}^{t_1} (f^0(x(t),t) + h^0(u(t),t))dt \tag{37}$$

rather than (21), where $K : X^n \to (X^k, \preccurlyeq)$ is convex and differentiable on X^n and f^0 and h^0 have the same properties as they have in (21), then the following sufficiency result is obtained :

<u>Theorem III.8.</u> If the optimal control problem is nonsingular and if there exists a measurable control $u^* : [t_0,t_1] \to \Omega$ satisfying the necessary conditions of the infimum principle, then u^* is globally superior (although not necessarily unique).

<u>Proof.</u> As in the proof of Theorem III.7, with $P_0^* = I$, and with the definitions (25)

and (26), the inequality (35) is obtained. Since necessarily

$$P^*(t_1) = \frac{\partial K}{\partial x}(x^*(t_1)) \quad , \tag{38}$$

equation (35) reads

$$x^0(t_1) - x^{0*}(t_1) + \frac{\partial K}{\partial x}(x^*(t_1))(x(t_1) - x^*(t_1)) \succcurlyeq 0 \quad . \tag{39}$$

Adding and subtracting identical terms in (39), viz. $K(x(t_1))$ and $K(x^*(t_1))$, respectively, (39) becomes

$$\begin{aligned} & x^0(t_1) + K(x(t_1)) - x^{0*}(t_1) - K(x^*(t_1)) - K(x(t_1)) + K(x^*(t_1)) \\ & + \frac{\partial K}{\partial x}(x^*(t_1))(x(t_1) - x^*(t_1)) \succcurlyeq 0 \quad . \end{aligned} \tag{40}$$

By the convexity of K, the sum of the last three terms in (40) is in the negative cone of $(X^k, \preccurlyeq)$. Thus,

$$J(u) - J(u^*) = x^0(t_1) + K(x(t_1)) - x^{0*}(t_1) - K(x^*(t_1)) \succcurlyeq 0$$

for all admissible u . (41)

IV. APPLICATIONS

In this section, the infimum principle is applied to nontrivial infimization problems. In Section IV.1, the Kalman-Bucy filter is rederived as the superior solution of a dynamic optimization problem with a matrix-valued cost criterion, viz. the error covariance matrix at some final time. In Section IV.2, an infimization problem is discussed, which is dual to the Kalman-Bucy filtering problem. Furthermore, the so-called separation theorem for a stochastic linear control problem with a quadratic cost functional is discussed. In Section IV.3, a superior solution to an uncertain optimal control problem with set-valued cost is found, where the uncertainty is described by set membership.

In this section, M_{nm} denotes the linear space of n by m matrices with real entries, M_n the abridged notation of M_{nn}, M'_n the linear space of n by n symmetric matrices with real entries, and $(M'_n, \preccurlyeq)$ the linear space M'_n partially ordered by positive-semidefinite differences (Example II.3 of [2]).

<u>Definition IV.1.</u> The linear operator $U : M_n \rightarrow M'_n$ is defined for all $A \in M_n$ by

$$UA = A + A' \quad . \tag{1}$$

<u>Definition IV.2.</u> The linear operator $T : M_{nm} \rightarrow M_{mn}$ is defined for all $A \in M_{nm}$ by

$$TA = A' \quad . \tag{2}$$

Obviously, by Definitions IV.1 and IV.2, U and T satisfy the operator equality

$$UT = TU = U \quad , \tag{3}$$

provided dimensions match.

IV.1. Rederivation of the Kalman-Bucy Filter

A slightly simplified version of the problem considered below has been presented by the authors in [10].

Statement of the Infimization Problem

Consider the n-th order linear time-varying dynamic system

$$\dot{x}(t) = A(t)x(t) + B(t)u(t) \tag{4}$$

$$x(t_0) = x_0 \quad . \tag{5}$$

The m-vector input u(.) is a white stochastic process with

$$E\{u(t)\} = 0 \text{ for all } t \tag{6}$$

$$E\{u(t)u'(s)\} = Q(t)\delta(t-s) \quad , \tag{7}$$

where $Q(t) \succcurlyeq 0 \in (M'_m, \preccurlyeq)$ for all t. The initial state x_0 is a random vector with

$$E\{x_0\} = \overline{x}_0 \tag{8}$$

$$E\{(x_0 - \overline{x}_0)(x_0 - \overline{x}_0)'\} = \Sigma_0 \succcurlyeq 0 \in (M'_n, \preccurlyeq) \quad . \tag{9}$$

The p-vector observation

$$y(t) = C(t)x(t) + v(t) \tag{10}$$

is corrupted by a white stochastic process v(.) with

$$E\{v(t)\} = 0 \text{ for all } t \tag{11}$$

$$E\{v(t)v'(s)\} = R(t)\delta(t-s) \quad , \tag{12}$$

where $R(t) \succnsim 0 \in (M'_p, \preccurlyeq)$ is a positive-definite matrix for all t. The noises u and v are assumed to be correlated according to

$$E\{u(t)v'(s)\} = S(t)\delta(t-s) \quad , \tag{13}$$

where $S(t) \in M_{mp}$ is such that the augmented matrix

$$\begin{bmatrix} Q(t) & S(t) \\ S'(t) & R(t) \end{bmatrix} \succ 0 \in (M_{(m+p)(m+p)}, \preccurlyeq) \tag{14}$$

is positive-semidefinite for all t. Furthermore, the random initial state x_0 is assumed to be independent of both $u(.)$ and $v(.)$.

It is desired to find an n-th order linear unbiased estimator with infimal conditional error covariance matrix at some given final time $t_1 > t_0$. In other words, the error $e(t)$ between the true state $x(t)$ and the state estimate $w(t)$ by the n-th order linear time-varying dynamic filter

$$\dot{w}(t) = F^*(t)w(t) + G^*(t)y(t) \tag{15}$$

$$w(t_0) = w_0^* \tag{16}$$

has to have conditional mean zero, i.e.,

$$E\{e(t) \mid y(s), t_0 \leq s \leq t\} = E\{x(t) - w(t) \mid y(s), t_0 \leq s \leq t\} = 0 \text{ for all } t \text{ ,} \tag{17}$$

and $F^* : [t_0, t_1] \to M_n$, $G^* : [t_0, t_1] \to M_{np}$, and $w_0^* \in R^n$ have to be found such that

$$\Sigma^*(t_1) = E\{e(t_1)e'(t_1) \mid y(s), t_0 \leq s \leq t\}\Big|_{F^*, G^*, w_0^*}$$

$$\preccurlyeq \Sigma(t_1) = E\{e(t_1)e'(t_1) \mid y(s), t_0 \leq s \leq t\}\Big|_{F, G, w_0} \tag{18}$$

for all other choices of F, G, and w_0.

For a discussion of the matrix-valued performance criterion, see [2, Remark IV.2].

Analysis of the Necessary Conditions of the Infimization Problem

For arbitrary choice of F, G, and w_0, the estimation error $e(t) = x(t) - w(t)$ satisfies the differential equation

$$\dot{e}(t) = (A(t) - F(t) - G(t)C(t))x(t) + F(t)e(t) + B(t)u(t) - G(t)v(t) \tag{19}$$

$$e(t_0) = x_0 - w_0 \text{ .} \tag{20}$$

Requiring unbiased estimates for all $t \in [t_0, t_1]$, (17), yields

$$w_0^* = \overline{x}_0 \tag{21}$$

and

$$F^*(t) = A(t) - G^*(t)C(t) \tag{22}$$

since u and v are zero-mean and $E\{x(t)\} \neq 0$ in general. Replacing F(t) by A(t) - G(t)C(t) in (19) and with (21), the error differential equation (19), (20) becomes

$$\dot{e}(t) = (A(t) - G(t)C(t))e(t) + B(t)u(t) - G(t)v(t) \tag{23}$$

$$e(t_0) = x(t_0) - \bar{x}_0 \ . \tag{24}$$

In order to obtain a deterministic optimal control problem, the error covariance matrix $\Sigma(t)$ is introduced

$$\Sigma(t) = E\{e(t)e'(t) \mid y(s),\ t_0 \leq s \leq t\} \ , \tag{25}$$

which satisfies the matrix differential equation

$$\dot{\Sigma}(t) = U(A(t) - G(t)C(t))\Sigma(t) + B(t)Q(t)B'(t) + G(t)R(t)G'(t) - UB(t)S(t)G'(t) \tag{26}$$

$$\Sigma(t_0) = \Sigma_0 \ . \tag{27}$$

In the remaining deterministic infimization problem, $\Sigma(t)$ is the (extended) state and G(t) is the control. The cost to be infimized is

$$J(G) = \Sigma(t_1) \ . \tag{28}$$

In this problem, the cost space is the entire state space (of dimension n(n+1)/2). The costate P(t) belongs to $\mathcal{L}((M'_n,\preccurlyeq),(M'_n,\preccurlyeq),\preccurlyeq)$ and the Hamiltonian is

$$H = P(t)\dot{\Sigma}(t) \ . \tag{29}$$

By Theorem II.7, if $G^* : [t_0,t_1] \rightarrow M_{np}$ is superior, then the following relations hold:

$$\begin{aligned}\dot{\Sigma}^*(t) = U(A(t) - G^*(t)C(t))\Sigma^*(t) + B(t)Q(t)B'(t)\\ + G^*(t)R(t)G^{*\prime}(t) - UB(t)S(t)G^{*\prime}(t)\end{aligned} \tag{30}$$

$$\Sigma^*(t_0) = \Sigma_0 \tag{31}$$

$$\dot{P}^*(t) = - P^*(t)U(A(t) - G^*(t)C(t)) \tag{32}$$

$$P^*(t_1) = P_0 \succcurlyeq 0 \in \mathcal{L}((M'_n,\preccurlyeq),(M'_n,\preccurlyeq),\preccurlyeq) \tag{33}$$

$$H(\Sigma^*(t),P^*(t),G^*(t),t) \preccurlyeq H(\Sigma^*(t),P^*(t),G(t),t)$$
$$\text{for all } t \in [t_0,t_1] \text{ and all } G(t) \in M_{np} \quad . \tag{34}$$

Since no singularity condition arises, P_0 in (33) could be taken to be the identity map by Remark II.6.

Observe, that the homogeneous differential equation in M'_n

$$\dot{\Sigma}(t) = U(A(t) - G(t)C(t))\Sigma(t) \tag{35}$$

$$\Sigma(t_0) = \Sigma_0 \tag{36}$$

has a positive-semidefinite solution $\Sigma(t)$ for all t, whenever Σ_0 is positive-semidefinite,

$$\Sigma(t) = \Phi_G(t,t_0)\Sigma_0 \quad , \tag{37}$$

where the transition operator $\Phi_G(.,.)$ is the solution of the operator differential equation

$$\frac{d}{dt}\Phi_G(t,t_0) = U(A(t) - G(t)C(t))\Phi_G(t,t_0) \tag{38}$$

$$\Phi_G(t_0,t_0) = I \quad . \tag{39}$$

Hence, the transition operator $\Phi_G(t,t_0) \in \mathcal{L}((M'_n,\preccurlyeq),(M'_n,\preccurlyeq),\preccurlyeq)$ is positive (Definition A.1) for all $t \in [t_0,t_1]$ and all possible choices of $G : [t_0,t_1] \to M_{np}$.

Now, since the solution of the costate differential equation (32), (33) is

$$P^*(t) = P_0\Phi_{G^*}(t_1,t) \quad , \tag{40}$$

it follows, that the costate $P^*(t)$ is positive for all t and all $P_0 \succ 0$. Therefore, the infimization of the Hamiltonian (29), (34) is achieved by infimizing $\dot{\Sigma}(t)$, hence,

$$U(A(t) - G^*(t)C(t))\Sigma^*(t) + G^*(t)R(t)G^{*\prime}(t) - UB(t)S(t)G^{*\prime}(t) + B(t)Q(t)B'(t)$$
$$\preccurlyeq U(A(t)-G(t)C(t))\Sigma^*(t) + G(t)R(t)G'(t) - UB(t)S(t)G'(t) + B(t)Q(t)B'(t)$$
$$\text{for all } t \in [t_0,t_1] \text{ and all } G(t) \in M_{np} \quad . \tag{41}$$

Since $\dot{\Sigma}(t)$ is quadratic in $G(t)$ with positive-hemidefinite ([2], Definition II.18) second Fréchet derivative (by the positive-definiteness of $R(t)$), it is necessary and sufficient,

that the first Fréchet derivative of $\dot{\Sigma}^*(t)$ with respect to $G(t)$ vanish at $G^*(t)$ in order that $\dot{\Sigma}^*(t)$ attain an infimum (by Theorems III.1, and III.2, and Remark III.3 of [2]). Thus, with (3)

$$\left.\frac{\partial \dot{\Sigma}^*(t)}{\partial G}\right|_{G = G^*(t)} = U(G^*(t)R(t) - \Sigma^*(t)C'(t) - B(t)S(t))T = 0 \quad , \tag{42}$$

which implies

$$G^*(t) = (\Sigma^*(t)C'(t) + B(t)S(t))R^{-1}(t) \quad . \tag{43}$$

<u>Solution to the Infimization Problem</u>

The superior n-th order linear unbiased filter for the plant (4), ... , (14) is

$$\dot{w}(t) = A(t)w(t) + (\Sigma(t)C'(t) + B(t)S(t))R^{-1}(t)(y(t) - C(t)w(t)) \tag{44}$$

$$w(t_0) = \bar{x}_0 \tag{45}$$

and the error covariance matrix $\Sigma(t) \in (M_n', \preccurlyeq)$ is precomputable from the matrix Riccati differential equation

$$\begin{aligned}\dot{\Sigma}(t) = {} & (A(t) - B(t)S(t)R^{-1}(t)C(t))\Sigma(t) + \Sigma(t)(A(t) - B(t)S(t)R^{-1}(t)C(t))' \\ & + B(t)Q(t)B'(t) - \Sigma(t)C'(t)R^{-1}(t)C(t)\Sigma(t) - B(t)S(t)R^{-1}(t)S'(t)B'(t)\end{aligned} \tag{46}$$

$$\Sigma(t_0) = \Sigma_0 \quad . \tag{47}$$

This result and the fact that the error covariance matrix $\Sigma(t_1)$ is infimized are well known[18], at least for $S(t) \equiv 0$. However, the derivation is new. The special case $p = m$ with $v(t) \equiv u(t)$ may be of some interest, if only for demonstration purposes, when only one noise generator is at hand. In this case, $R(t) = Q(t) = S(t) = S'(t)$ for all t.

<u>Sufficiency Analysis</u>

In the sequel it is proved, that the Kalman-Bucy filter indeed is a globally superior solution to the infimization problem. If in (13), the matrix $S(t)$ is set identically equal zero for the sake of simpler arithmetics, then, the deterministic Kalman-Bucy filtering problem consists

of the dynamic system

$$\dot{\Sigma}(t) = (A(t) - G(t)C(t))\Sigma(t) + \Sigma(t)(A(t) - G(t)C(t))' \\ + B(t)Q(t)B'(t) + G(t)R(t)G'(t) \tag{48}$$

$$\Sigma(t_0) = \Sigma_0 \tag{49}$$

and the matrix-valued cost

$$J(G) = \Sigma(t_1) \quad . \tag{50}$$

In the context of the Hamilton-Jacobi-Bellman theory of Section III.1, this problem is interpreted as follows : In (3) of Section III

$$K(\Sigma(t_1),t_1) = \Sigma(t_1) \tag{51}$$

$$L(\Sigma(t),G(t),t) \equiv 0 \quad . \tag{52}$$

The target set is $S = (M'_n,\preccurlyeq) \times \{t_1\}$, the regions $Z = (M'_n,\preccurlyeq) \times \{t \in R \mid t \leq t_1\}$ and Π = (positive cone of $\mathcal{L}((M'_n,\preccurlyeq),(M'_n,\preccurlyeq),\preccurlyeq) \times \{t \in R \mid t \leq t_1\}$. The Hamiltonian (4) of Section III is

$$H(\Sigma,P,G,t) = P((A-GC)\Sigma + \Sigma(A-GC)' + BQB' + GRG') \tag{53}$$

where the notation of the time dependence is suppressed. Since the costate $P(t)$ with boundary condition $P(t_1) = \frac{\partial K}{\partial \Sigma}(\Sigma(t_1),t_1) = I$ necessarily is a positive map for all $t \leq t_1$, the H-infimal control is

$$G^0(\Sigma,P,t) = \Sigma C'(t)R^{-1}(t) \quad . \tag{54}$$

With (54), (48) and (49) become

$$\dot{\Sigma}(t) = A(t)\Sigma(t) + \Sigma(t)A'(t) + B(t)Q(t)B'(t) - \Sigma(t)C'(t)R^{-1}(t)C(t)\Sigma(t) \tag{55}$$

$$\Sigma(t_0) = \Sigma_0 \quad . \tag{56}$$

After defining the 2n by 2n matrix

$$\begin{bmatrix} A(t) & B(t)Q(t)B'(t) \\ C'(t)R^{-1}(t)C(t) & -A'(t) \end{bmatrix} \tag{57}$$

and its 2n by 2n transition matrix

$$\begin{bmatrix} \Psi_{11}(t,t_0) & \Psi_{12}(t,t_0) \\ \Psi_{21}(t,t_0) & \Psi_{22}(t,t_0) \end{bmatrix} , \tag{58}$$

the closed-form solution of (55), (56) can be written as

$$\Sigma(t) = (\Psi_{11}(t,t_0)\Sigma_0 + \Psi_{12}(t,t_0))(\Psi_{22}(t,t_0) + \Psi_{21}(t,t_0)\Sigma_0)^{-1} . \tag{59}$$

In the context of the Hamilton-Jacobi-Bellman theory, where the initial state is denoted by Σ and the initial time by t, the cost (50) then is

$$J(G) = J(\Sigma,t) = (\Psi_{11}(t_1,t)\Sigma + \Psi_{12}(t_1,t))(\Psi_{22}(t_1,t) + \Psi_{21}(t_1,t)\Sigma)^{-1} . \tag{60}$$

Clearly, $J(\Sigma,t)$ of (60) satisfies the boundary condition (8) of Section III on S,

$$J(\Sigma,t_1) = \Sigma = K(\Sigma,t_1) , \tag{61}$$

since $\Psi_{11}(t_1,t_1) = \Psi_{22}(t_1,t_1) = I$ and $\Psi_{12}(t_1,t_1) = \Psi_{21}(t_1,t_1) = 0$. In order to verify the Hamilton-Jacobi-Bellman partial differential equation

$$\frac{\partial J}{\partial t}(\Sigma,t) + H(\Sigma,\frac{\partial J}{\partial \Sigma}(\Sigma,t), G^0(\Sigma,\frac{\partial J}{\partial \Sigma}(\Sigma,t),t),t) = 0 , \tag{62}$$

the derivatives $\frac{\partial J}{\partial t}(\Sigma,t) \in \mathcal{L}(R,M_n')$ (which is isomorphic to M_n') and $\frac{\partial J}{\partial \Sigma}(\Sigma,t) \in \mathcal{L}(M_n',M_n')$ are calculated :

$$\begin{aligned} \frac{\partial J}{\partial t}(\Sigma,t) = \; & (-\Psi_{11}(t_1,t)A\Sigma - \Psi_{12}(t_1,t)C'R^{-1}C\Sigma - \Psi_{11}(t_1,t)BQB' + \Psi_{12}(t_1,t)A') \\ & \times (\Psi_{22}(t_1,t) + \Psi_{21}(t_1,t)\Sigma)^{-1} \\ & - (\Psi_{11}(t_1,t)\Sigma + \Psi_{12}(t_1,t))(\Psi_{22}(t_1,t) + \Psi_{21}(t_1,t)\Sigma)^{-1} \\ & \times (-\Psi_{21}(t_1,t)BQB' + \Psi_{22}(t_1,t)A' - \Psi_{21}(t_1,t)A\Sigma - \Psi_{22}(t_1,t)C'R^{-1}C\Sigma) \\ & \times (\Psi_{22}(t_1,t) + \Psi_{21}(t_1,t)\Sigma)^{-1} \end{aligned} \tag{63}$$

$$\begin{aligned} \frac{\partial J}{\partial \Sigma}(\Sigma,t) = \; & \Psi_{11}(t_1,t)T(\Psi_{22}(t_1,t) + \Psi_{21}(t_1,t)\Sigma)'^{-1}T \\ & - (\Psi_{11}(t_1,t)\Sigma + \Psi_{12}(t_1,t))(\Psi_{22}(t_1,t) + \Psi_{21}(t_1,t)\Sigma)^{-1} \\ & \times \Psi_{21}(t_1,t)T(\Psi_{22}(t_1,t) + \Psi_{21}(t_1,t)\Sigma)'^{-1}T \end{aligned} \tag{64}$$

where in (63) and (64) (and below) the matrices A, B, C, Q, and R are understood to be evaluated at time t, and T is the transposition operator (Definition IV.2).

Combining (53), (54), (63), and (64), it is easily seen that the Hamilton-Jacobi-Bellman equation (62) is satisfied, since all terms cancel. Hence, by Theorem III.3 and Corollaries III.5 and III.6, the Kalman-Bucy filter is a globally superior solution to the dynamic infimization problem (48), (49), (50), since the costate P(t) is necessarily a positive map.

Clearly, the above analysis gives the same result for the more general case of $S(t) \not\equiv 0$. Then, the matrix (57) is

$$\begin{bmatrix} A(t) - B(t)S(t)R^{-1}(t)C(t) & B(t)(Q(t) - S(t)R^{-1}(t)S'(t))B'(t) \\ C'(t)R^{-1}(t)C(t) & -A'(t) + C'(t)R^{-1}(t)S'(t)B'(t) \end{bmatrix} \tag{65}$$

and the analysis goes through in an analogous fashion.

IV.2. Related Infimization Problems

The Dual Infimization Problem to the Kalman-Bucy Filtering Problem

The deterministic linear quadratic regulator problem and the least-squares filtering problem for the linear system (4) are often called dual to each other. In the context of non-scalar-valued performance criteria, the question arises, in what sense the linear quadratic regulator problem is dual to the infimization problem of Section IV.1 with its matrix-valued performance criterion (18).

The control $u^* : [t_0, t_1] \to R^m$ for the n-th order linear system

$$\dot{x}(t) = A(t)x(t) + B(t)u(t) \tag{66}$$

$$x(t_0) = x_0 \ , \tag{67}$$

which is optimal with respect to the cost functional

$$J(u) = x'(t_1)Fx(t_1) + \int_{t_0}^{t_1} (x'(t)Q(t)x(t) + u'(t)R(t)u(t))dt$$

$$F \succcurlyeq 0 \in (M'_n, \preccurlyeq), \quad Q(t) \succcurlyeq 0 \in (M'_n, \preccurlyeq), \quad \text{and } R(t) \succcurlyeq\!\!\!\!/\; 0 \in (M'_m, \preccurlyeq)$$

$$\text{(i.e., positive-definite) for all } t \in [t_0, t_1] \tag{68}$$

is superior in the sense, that it is optimal for every arbitrary initial state $x_0 \in R^n$. In other words, u^* infimizes the function-valued cost $J : L_2^m(t_0, t_1) \rightarrow (C(R^n), \preccurlyeq)$ ([19, p.2, Example 3, i.e., $C(R^n)$ partially ordered by the pointwise total order of $(R,<)$]),

$$J(u) = x'(t_1, x_0)Fx(t_1, x_0) + \int_{t_0}^{t_1} (x'(t_1, x_0)Q(t)x(t_1, x_0) + u'(t)R(t)u(t))dt \tag{69}$$

The easy verification by the results of Sections II and III is left to the reader (see [9, p. 83]).

The Separation Theorem

An interesting exercise for the application of the infimum principle is provided by the so-called separation theorem for the stochastic linear quadratic Gaussian regulator problem [9, Section V.3]. Since the separation of the estimator from the regulator is implemented in the constraints of the infimization problem, the analysis of this exercise clearly does not prove the separation theorem; it merely verifies the fact, that the quadratic cost functional S and the estimation error covariance matrix $\Sigma(t_1)$ are infimized simultaneously. Thus, the compounded non-scalar-valued cost to be infimized is

$$J = (S, \Sigma(t_1)) \in ((R, \leqslant) \times (M'_n, \preccurlyeq), \preccurlyeq) \quad , \tag{70}$$

where the partial order of $R \times M'_n$ is defined by

$$J_1 = (S_1, \Sigma_1) \preccurlyeq J_2 = (S_2, \Sigma_2) \Longleftrightarrow S_1 \leqslant S_2 \text{ in } (R, \leqslant) \text{ and } \Sigma_1 \preccurlyeq \Sigma_2 \text{ in } (M'_n, \preccurlyeq). \tag{71}$$

IV.3. An Example with a Set-Valued Cost

In this section, an optimization problem with a set-valued cost is investigated. A superior solution in the sense of the partial order of set inclusion [2, Example II.4] is found. The problem involves a linear uncertain dynamic system, where the uncertainty is described by set-membership.

Statement of the Infimization Problem

Consider the observed uncertain n-th order linear dynamic system

$$\dot{x}(t) = A(t)x(t) + B(t)(u(t) + w(t)) \tag{72}$$

$$x(t_0) = x_0 \tag{73}$$

$$y(t) = C(t)x(t) + v(t) \quad . \tag{74}$$

$B(t)$ is an n by m matrix of full rank $m \leqslant n$, $A(t)$ is n by n, $C(t)$ is r by n, and they all are piecewise continuous on $[t_0,t_1]$. The initial state $x_0 \in R^n$, the control uncertainty $w(.) \in L_2^m(t_0,t_1)$, and the observation uncertainty $v(.) \in L_2^r(t_0,t_1)$ are merely known to belong to a set, i.e., $(x_0,w(.),v(.)) \in S \subset R^n \times L_2^m(t_0,t_1) \times L_2^r(t_0,t_1)$. The convex constraint set S is described by its support functional [17] $\sigma(q_1,q_2(.),q_3(.);S)$, $q_1 \in R^n$, $q_2(.) \in L_2^m(t_0,t_1)$, $q_3(.) \in L_2^r(t_0,t_1)$,

$$\sigma(q_1,q_2(.),q_3(.);S) = \left(q_1{}'Fq_1 + \int_{t_0}^{t_1} (q_2'(t)Q(t)q_2(t) + q_3'(t)R(t)q_3(t))dt\right)^{1/2} ,$$

where $F \succcurlyeq 0 \in (M_n', \preccurlyeq)$, $Q(t) \succcurlyeq 0 \in (M_m', \preccurlyeq)$ for all $t \in [t_0,t_1]$, and

$R(t) \succ 0 \in (M_r', \prec)$ (i.e., positive-definite) for all $t \in [t_0,t_1]$. (75)

In the special case, where F and Q(t) are invertible, this set-membership constraint of the uncertainty is equivalent to the "energy constraint"

$$x_0'F^{-1}x_0 + \int_{t_0}^{t_1} (w'(t)Q(t)w(t) + v'(t)R(t)v(t))dt \leqslant 1 \quad . \tag{76}$$

For $t \in [t_0, t_1]$ find an m by r matrix $N^*(t)$ such that the linear output feedback control law

$$u(t) = N^*(t)y(t) \tag{77}$$

generates a set $\{x(t_1)\}_{N^*}$ of uncertain final states $x(t_1)$, which is infimal in the sense of set inclusion, i.e.,

$$J(N^*) = \{x(t_1)\}_{N^*} \subseteq J(N) = \{x(t_1)\}_N \text{ for all } N(.) \quad . \tag{78}$$

Analysis of the Infimization Problem

With the linear output feedback control law $u(t) = N(t)y(t)$, (72), (73), and (74) combine to

$$\dot{x}(t) = (A(t) + B(t)N(t)C(t))x(t) + B(t)(w(t) + N(t)v(t)) \tag{79}$$

$$x(t_0) = x_0 \quad . \tag{80}$$

For any particular uncertainty x_0, $w(.)$, $v(.)$, the uncertain final state can be written as

$$x(t_1) = \Phi(t_1,t_0)x_0 + \int_{t_0}^{t_1} \Phi(t_1,t)B(t)w(t)dt + \int_{t_0}^{t_1} \Phi(t_1,t)B(t)N(t)v(t)dt \quad , \tag{81}$$

where $\Phi(.,.)$ is the state transition matrix associated with the systems matrix $A(t) + B(t)N(t)C(t)$, or

$$x(t_1) = L(x_0,w,v) \quad , \tag{82}$$

where $L : R^n \times L_2^m(t_0,t_1) \times L_2^r(t_0,t_1) \rightarrow R^n$ is the linear operator defined by (81). With the usual inner product on $R^n \times L_2^m(t_0,t_1) \times L_2^r(t_0,t_1)$ and with the adjoint operator L^* of L, the support functional $\sigma(q;\{x(t_1)\}_N)$, $q \in R^n$, can easily be written as [17]

$$\begin{aligned} \sigma(q;\{x(t_1)\}_N) &= \sigma(q;L\{(x_0,w,v)\}) = \sigma(L^*q;S) \\ &= \Big(q' \Big(\Phi(t_1,t_0)F\Phi'(t_1,t_0) + \int_{t_0}^{t_1} \Phi(t_1,t)B(t)Q(t)B'(t)\Phi'(t_1,t)dt \\ &\qquad + \int_{t_0}^{t_1} \Phi(t_1,t)B(t)N(t)R(t)N'(t)B'(t)\Phi'(t_1,t)dt\Big) q\Big)^{1/2} \\ &= (q'M(t_1)_N q)^{1/2} \quad . \end{aligned} \tag{83}$$

Clearly, $M(t_1)_N \in (M'_n,\preccurlyeq)$ defined in (83) is positive-semidefinite. Furthermore, the following partial order inequalities are equivalent :

$$\{x(t_1)\}_{N*} \subseteq \{x(t_1)\}_N \qquad \text{(partial order of set-inclusion [2, Example II.4])}$$

$$\iff \sigma(q;\{x(t_1)\}_{N*}) \preccurlyeq \sigma(q;\{x(t_1)\}_N) \qquad \text{(partial order of pointwise total order of } (R,\leqslant) \text{ [19, p. 2, Example 3])}$$

$$\iff M(t_1)_{N*} \preccurlyeq M(t_1)_N \quad . \qquad \text{(partial order of positive-semidefinite differences [2, Example II.3])} \tag{84}$$

Thus, the infimization problem at hand reduces to the following infimization problem : For

$$\dot{M}(t) = (A(t) + B(t)N(t)C(t))M(t) + M(t)(A(t) + B(t)N(t)C(t))'$$

$$+ B(t)Q(t)B'(t) + B(t)N(t)R(t)N'(t)B'(t) \tag{85}$$

$$M(t_0) = F \quad , \tag{86}$$

find $N : [t_0,t_1] \to M_{mr}$, such that

$$M(t_1) \text{ is infimal with respect to } N(.) \ . \tag{87}$$

This infimization problem is quite analogous to the infimization problem (26), (27), (28) in the analysis of the Kalman-Bucy filter. Here, the state is $M(t) \in M'_n$ and the control is $N(t) \in M_{mr}$. The cost space $(M'_n,\preccurlyeq)$ is the entire state space. As in (29) through (41), the costate $P(t) \in \mathcal{L}((M'_n,\preccurlyeq),(M'_n,\preccurlyeq),\preccurlyeq)$ turns out to be a positive map for all $t \in [t_0,t_1]$. Therefore, the infimization with respect to $N(t)$ of the Hamiltonian

$$H = P(t)\dot{M}(t) \tag{88}$$

reduces to the infimization of $\dot{M}(t)$ with respect to $N(t)$. For the latter, by Theorems III.1, and III.2, and Remark III.3 of [2], the following conditions are necessary and sufficient

$$\frac{\partial \dot{M}}{\partial N} = U(M(t)C'(t) + B(t)N(t)R(t))TB(t) = 0 \tag{89}$$

$$d^2M(N,d^2N) = B(t)dNR(t)dN'B'(t) \succ 0 \text{ for all } dN \neq 0 \tag{90}$$

where U and T have been defined in the Definitions IV.1 and IV.2. Equation (89) is

satisfied by any matrix $N^*(t)$ satisfying

$$B(t)N^*(t) = - M(t)C'(t)R^{-1}(t) \quad . \tag{91}$$

Since $B(t)$ is of full rank $m \leq n$, (90) is satisfied, and (91) can be written with the pseudoinverse $B^{\dagger}(t)$ of $B(t)$ as

$$N^*(t) = - B^{\dagger}(t)M(t)C'(t)R^{-1}(t) = - (B'(t)B(t))^{-1}B'(t)M(t)C'(t)R^{-1}(t) \quad . \tag{92}$$

As in the case of the Kalman-Bucy filter, the Hamilton-Bellman-Jacobi theory of Section III.1 verifies, that $N^*(t)$ of (91) and (92) is a globally superior feedback matrix. The infimal ellipsoidal set $\{x(t_1)\}_{N^*}$ of all possible final states is described by the support functional

$$\sigma(q;\{x(t_1)\}_{N^*}) = (q'M^*(t_1)q)^{1/2} \quad , \tag{93}$$

where $M^*(t_1)$ can be calculated from the matrix Riccati differential equation

$$\dot{M}^*(t) = A(t)M^*(t) + M^*(t)A'(t) + B(t)Q(t)B'(t)$$
$$- M^*(t)C'(t)R^{-1}(t)C(t)M^*(t) \tag{94}$$

$$M^*(t_0) = F \quad . \tag{95}$$

V. CONCLUSIONS

This research has been concerned with superior solutions to optimal control problems with non-scalar-valued performance criteria. The main contribution to the theory of optimal control is the infimum principle in Section II, which constitutes necessary conditions for a control to be superior with respect to a non-scalar-valued performance criterion attaining its value in a finite-dimensional abstract partially ordered cost space, the positive cone of which is closed and has nonempty interior. Further contributions include the sufficiency results for a control to be superior in Section III, most notably the extension of the Hamilton-Jacobi-Bellman theory to the case of non-scalar-valued performance criteria. In Section IV, the applicability of the theory of Sections II and III has been demonstrated with two non-trivial examples, viz. the Kalman-Bucy filter and an uncertain optimal control problem with set-membership description of the uncertainty.

In order to keep the proofs simple, the cost spaces have been restricted to be finite-dimensional. Conceptually, the theory of this paper can be extended to infinite-dimensional cost spaces, provided their positive cones are endowed with suitable topological properties. The additional properties needed have been discussed in [16], where a theory of superior solutions to problems of mathematical programming in abstract infinite-dimensional partially ordered spaces has been published.

In the theory of optimal control for non-scalar-valued performance criteria, the problem of existence of superior solutions still is open to further research. Presently, it is unclear what types of reasonable assumptions are needed in the problem statement, in order to obtain existence of superior solutions rather than existence of noniferior solutions. The major difficulty in the existence problem has been described in [2, Example IV.5 and Section V, last paragraph].

APPENDIX

Definition A.1. A map f from the linear poset $(X,\preccurlyeq)$ into the linear poset $(Y,\preccurlyeq)$ is called order-positive, if $x \succcurlyeq 0 \in (X,\preccurlyeq)$ implies $f(x) \succcurlyeq 0 \in (Y,\preccurlyeq)$. Furthermore, $\mathcal{L}((X,\preccurlyeq),(Y,\preccurlyeq),\preccurlyeq)$ denotes the linear poset of all linear maps from X into Y, which is partially ordered by the cone of all positive linear maps.

Lemma A.2. The positive cone of the directed linear poset $(X,\preccurlyeq)$ is the supremal [2, Definition II.5] set (in the sense of the partial order of set-inclusion) of the subsets of $(X,\preccurlyeq)$, which are mapped into the positive cone of the linear poset $(Y,\preccurlyeq)$ by all positive linear maps.

REFERENCES

1. M. Athans and P.L. Falb, "Optimal Control", McGraw-Hill Book Co, New York, N.Y., 1966.
2. M. Athans and H.P. Geering, "Necessary and Sufficient Conditions for Differentiable Non-Scalar-Valued Functions to Attain Extrema", IEEE Trans. on Automatic Control, AC-18(1973), April issue.
3. R. Bellman (ed.), "Mathematical Optimization Techniques", University of California Press, Berkeley, Calif., 1963.
4. G. Birkhoff, "Lattice Theory", Coll. Publ. 25, 3rd ed., American Mathematical Society, Providence, R.I., 1967.
5. L. Cesari, "Existence Theorems for Weak and Usual Optimal Solutions in Lagrange Problems with Unilateral Constraints", Trans. Amer. Math. Soc., 124(1966), 369-412.
6. S.S.L. Chang, "General Theory of Optimal Processes", SIAM J. Control, 4(1966), 46-55.
7. D.H. Chyung, "Optimal Systems with Multiple Cost Functionals", SIAM J. Control, 5(1967), 345-351.
8. N.O. DaCunha and E. Polak, "Constrained Minimization under Vector-Valued Criteria in Finite Dimensional Spaces", J. Math. Anal. Appl., 19(1967), 103-124.
9. H.P. Geering, "Optimal Control Theory for Non-Scalar-Valued Performance Criteria", Ph.D. Dissertation, Massachusetts Institute of Technology, Department of Electrical Engineering, Cambridge, Massachusetts, August, 1971.

10. H.P. Geering and M. Athans, "Optimal Control Theory for Non-Scalar-Valued Performance Criteria", Proc. of the Fifth Annual Princeton Conference on Information Sciences and Systems, Princeton, N.J., March, 1971.

11. H. Halkin, "On the Necessary Condition for Optimal Control of Non-Linear Systems", J. Analyse Mathématique, 12(1964), 1-82.

12. H. Halkin, "Topological Aspects of Optimal Control of Dynamical Polysystems", Contributions to Differential Equations, 3(1964), 377-385.

13. E.B. Lee and L. Markus, "Foundations of Optimal Control Theory", J. Wiley and Sons, Inc., New York, N.Y., 1967.

14. L.W. Neustadt, "A General Theory of Extremals", J. of Computer Sciences and Systems, 3(1969), 57-92.

15. C. Olech, "Existence Theorems for Optimal Problems with Vector Valued Cost Function", Technical Report 67-6, Center of Dynamical Systems, Brown University, Providence, R.I., 1967.

16. K. Ritter, "Optimization Theory in Linear Spaces", Math. Ann., Part I: 182(1969), 189-206, Part II: 183(1969), 169-180, Part III: 184(1970), 133-154.

17. R.T. Rockafellar, "Convex Analysis", Princeton University Press, Princeton, N.J.,1970.

18. E. Tse, "On the Optimal Control of Linear Systems with Incomplete Information", Ph.D. Dissertation, Massachusetts Institute of Technology, Department of Electrical Engineering, Cambridge, Mass., November, 1969.

19. B.Z. Vulikh, "Introduction to the Theory of Partially Ordered Spaces", Wolters-Noordhoff Sc. Publ. Co., Groningen, Netherlands, 1967.

20. H.S. Witsenhausen, "Minimax Control of Uncertain Systems", Report ESL-R-269, Electronic Systems Laboratory, Massachusetts Institute of Technology, Cambridge, Mass., 1966.

21. L.A. Zadeh, "Optimality and Non-Scalar-Valued Performance Criteria", IEEE Trans. on Automatic Control, AC-8(1963), 59-60.